The Predictive Power of Financial Blogs

Ben Frisbee

Haverford College
Advisor: Professor Ghosh
Senior Economics Thesis

April 29, 2010

Abstract

This paper examines the relationships between the investor sentiment of various reputable financial bloggers and market movements. Bullish, bearish and neutral blogger sentiment percentages are taken from the weekly “Blogger Sentiment Poll” of the financial blog Ticker Sense and compared with changes in market prices and market volume of the S&P 500 between the dates of July 10, 2006 and December 21, 2009. The bloggers are very inaccurate, with increases in bullishness raising the probability of market bearishness and increases in bearishness raising the probability of market bullishness over two-week, one-month and 3-month time periods. Disagreement in blogger sentiment is not associated with increased trading volume.
# Table of Contents

1 Introduction______________________________________________________3

2 Literature Review_________________________________________________5

3 *Ticker Sense*______________________________________________________8

4 Data___________________________________________________________10

5 Results and Analysis______________________________________________14

6 Conclusion______________________________________________________24

7 Appendix_______________________________________________________26

8 References______________________________________________________28
1 Introduction

This paper examines the correlations between market predictions made by various financial bloggers and the movements of market prices and volume. Armed with data from the well-known “Blogger Sentiment Poll” of the financial blog Ticker Sense (www.TickerSense.com), this paper attempts to answer questions of how accurate financial bloggers are, what factors might influence their predictions, and whether we can gain any information about future market directions and volume changes from their forecasts. Because these accumulated predictions are really measures of investor opinion, the answers to these questions can also be useful in understanding the effects that general investor sentiment has on the market.

Studies exploring the accuracies of financial “experts,” generally in the form of analysts, economists or professional money managers have been varied in scope and mixed in results, but more often than not call into question the experts’ abilities as market forecasters. Studies on the effects that changes in investor sentiment have on market price movements and volume generally find that investor sentiment can help explain market volatility, but does not seem to have a defined effect on either price or volume movements.¹

However, studies in this area often find differing results, leaving clear conclusions temporarily unattainable. This paper plans to add to the research of expert financial opinion and investor sentiment in the vein of this relatively new phenomenon of financial blogs. Even though they are a fairly recent concept, financial blogs are rising in number, probably due to the fact that they can be a great way to disseminate financial information, voice opinion or advertise your

own analytical tools and financial prowess. Financial blogs are also growing in popularity among investors due to an increasing body of more sophisticated and insightful bloggers, often making these information hubs a regular part of daily investor research. The impact that this relatively new creation has on a broad range of investors makes it an interesting area in which to continue this type of research.
2 Literature Review

The first series of papers that helped inspire my research provide complex analysis on the effects that investor sentiment, macroeconomic policy or the media have on the market. Several of these papers dissect the semantic orientation of written financial content (such as message board posts or macroeconomic policy updates) using linguistics methods, and then see if changes in the content over time affect various market variables.

Antweiler and Frank [2001] analyze the power that bullish, bearish or neutral internet stock message board posts have in predicting future market movements and volatility, as well as whether investor agreement or disagreement in these posts leads to greater or less trading volume. Using automated linguistics programs to compute the bullishness/bearishness of over 1.5 million posts on the message boards of Yahoo! Finance and Raging Bull, they concluded that these messages help predict market volatility, but not actual up-or-down market movements. Also, they found that a greater disagreement between the messages did not see an increase in trading volume, but agreement among the posts was correlated with lower trade volumes.

Baker and Wurgler [2007] produced a paper looking at the effects that investor sentiment had on stock performance. Their investor sentiment data was constructed using different combinations of proxies, such as trading volume, dividend premium, and first-day returns on IPOs. They found that stocks that were more speculative and difficult to arbitrage were affected by investor sentiment more than safer, easy-to-arbitrage stocks. They also found that the average returns of safer, more bond-like stocks were greater during periods of high investor sentiment than the average returns of riskier, more speculative stocks. This finding goes against the
theoretical model that riskier assets must have higher returns to compensate the owners for burdening that risk.\textsuperscript{2}

Lucca and Trebbi [2009] researched changes in short and long-term Treasury yields based on linguistic changes in the content of Federal Open Market Committee (FOMC) statements. FOMC statements are released every three months or so, and contain information on the Fed’s current interest rate policy and future outlook. Using their own complex new linguistics methods, they rated each statement on a hawkish (higher interest rates) or dovish (lower interest rates) scale, and used high-frequency Treasury yield data to analyze what kinds of effects the language of the statements had on the yields. Lucca and Trebbi found that while short-term Treasury yield rates responded mostly to the actual changes in policy rates, longer-term Treasury yields were affected more by the (often subtle) changes in the language of the policy statements.

Paul Tetlock [2003] used linguistic software packages to break down the daily content from a popular \textit{Wall Street Journal} column into bullish, bearish and neutral information in order to analyze the effects that media sentiment has on the market. Tetlock found that high media pessimism was correlated with downward movements in market prices, and that abnormally high or low pessimism led to greater market trading volume.

The second set of papers that helped me refine my questions and narrow my research includes more straight-forward investigations into the accuracies of various financial experts and the impact that these experts have on the market.

Alfred Cowles [1932] called financial expertise in stock-picking into question. This paper found that financial companies’ recommendations often did worse than the general market, that recommendations made at random would have fared better, and that the best individual analyst performances were most likely due to chance.

Kent Womack [1996] found significant investment value in various brokerage analyst recommendations by looking at the value-added benefit post-recommendation. While buy recommendations exhibited a smaller and more short-lived (but still significant) return, sell recommendations had a stronger magnitude in their negative impact and were significant for longer periods of time. This paper found definite predictive power in these analysts’ predictions.

Adam Shapiro [2006] examined the stock return implications of Jim Cramer’s stock picks in the CNBC show *Mad Money*. Shapiro found that Cramer’s recommendations did indeed have an abnormally high return – both in the overnight performance of the individual stocks and in the longer term performances of portfolios acting upon his picks. Whether because of investor frenzy or because he was actually right on his own, Shapiro finds that Cramer is one “expert” who backs up his analysis with real performance.

My paper looks to bridge these various research papers in a new vein of financial information – financial blogs. Since financial blogs are a relatively new concept little research has been done on them. By using the data from the blogger sentiment poll, my paper can provide information about the accuracies of blogger predictions as well as the predictive ability that investor sentiment has on market prices and volume levels.
3 **Ticker Sense**

Ticker Sense is a well-respected financial blog run by Birinyi Associates, a financial research and money management firm located in Westport, Connecticut. Founded by company analysts in 2005, the blog discusses most anything finance-related, with a focus on U.S. stocks. Ticker Sense has been ranked one of the top 25 financial blogs by the well-known blog 24/7 Wall Street three times, in 2006, 2007, and 2009 (the 2009 article was printed in Time Magazine). Website traffic data from web information company Alexa ranks Ticker Sense as the 466,001st most popular website in the United States, with 235 other sites linking to it.

Blog posts are made generally once or twice a day with original market analysis supplemented by charts, tables and graphs. Recent posts (at the time of writing) include a composite chart of the price movements of the S&P 500 the day after a price decline greater than 2% since March of 2009, advice and reactions from earnings releases, breakdowns of company analysts’ sector ratings, and of course, the weekly “Blogger Sentiment Poll.”

Ticker Sense has this to say about the poll:

The Ticker Sense Blogger Sentiment Poll is a survey of the web's most prominent investment bloggers, asking "What is your outlook on the S&P 500 for the next 30 days?" Conducted on a weekly basis, the poll is sent to participants each Thursday, and the results are released on Ticker Sense each Monday. The goal of this poll is to gain a consensus view on the market from the top investment bloggers -- a community that continues to grow as a valued source of investment insight. © Copyright 2009 Ticker Sense Blogger Sentiment Poll.

Poll results are displayed in a pie chart with percentages of bullish, bearish and neutral bloggers, along with a line graph of historical bullish and bearish percentages. The poll generally

---

gets anywhere from 10-30 responses from its community of bloggers. The individual picks by each blog are shown at the end of each post. A short sentence of commentary sums up the results of the poll, such as, “Bloggers remained bullish following February’s gain.” Figure 1 gives an example of the information present in each poll post, this one posted on March 8th, 2010.

**Figure 1**

The poll remains moderately bullish this week.

Other financial bloggers have studied the Blogger Sentiment Poll as well. *Trader’s Narrative* examined periods of extreme swings in blogger sentiment, finding that the poll predictions were especially useless during these times. Other analysis, including a review done by financial blog *Zignals*, breaks down the accuracy of each individual poll participant over time.

---

4 Data

The data used in this paper consists of various stock market information from the S&P 500 for each poll observation (every day the Blogger Sentiment Poll was released since its inception on July 10\textsuperscript{th}, 2006 through the end of 2009, 173 observations in total), as well as the results of each poll. The S&P 500 is the market that suits this research the most since it is the market that the blog refers to when asking for predictions. For each observation date, 1-day, 5-day, 10-day, 20-day, and 60-day growth rates (in percentage values) are constructed for the opening, high, low, and closing market prices and volume from that date. (For example, “open1” is the set of 1-day growth rates of the opening market prices for each poll date, close5 is the set of 5-day growth rates of the closing market prices for each poll date, “volume60” is the set of 60-day growth rates of trading volume for each poll date, etc.) Since the S&P 500 is generally open 5 days per week, these growth rates refer to 1-day, 1-week, 2-weeks, 1-month, and 3-month growth rates. These statistics were all constructed using data from Yahoo! Finance. Bullish, bearish and neutral percentages were collected from each poll in the date range posted on Ticker Sense. Finally, the federal funds rate was obtained for each poll date from FederalReserve.gov.

In all regressions, the federal funds rate was used as a control variable to explain the day-to-day movements of the market. The funds rate is a powerful indicator of market activity, as it affects the rate at which investors discount the future payments of assets (such as dividends or growth of a stock). It is also a macroeconomic tool used by the Federal Reserve to tighten or loosen the flow of money. Because of this, it can be seen as a present and future indicator of macroeconomic and monetary policy. Its broad market and economic implications allow the funds rate to do the job of controlling for normal day to day market movements and information.
Variables constructed from the data (besides the growth rates) include “Diff,” a variable measuring the level of differing sentiment in the poll, equaling the sum of the bullish and bearish percentages less the neutral percentage for each poll. This way, the variable increases only when bullish and bearish sentiment increase together, not just one along with neutral, effectively capturing when the polled bloggers disagree on market outlook.

The other constructed variables were three dummy variables for each type (open, high, low and close) and range (1-day, 5-day, etc.) of price growth rates, signifying if a particular growth rate from a particular blog post was bullish, bearish, or neutral. For example, the variable “open1bu” takes a value of 1 if a specific 1-day open growth rate falls into the bullish range for the 1-day open growth rate values, and a 0 if it falls into either the neutral or bearish range. Bullish, bearish and neutral ranges were created by subtracting and adding the standard deviation of these price growth rate variables multiplied by 0.430727 to the mean, effectively splitting the normal distribution of these variables into thirds. While not perfectly normal, the price growth rates are normal enough for this procedure to systematically produce a balanced number of bearish, neutral and bullish growth rates for each of these variables.

Each variable was checked for unit roots using the Dickey-Fuller test at the 10% level. The only variable with a p-value above this level was the federal funds rate, which, due to its use in the regressions as a dependent control variable, was not found to be a concern. Table 1 gives the mean, median, standard deviation, minimum and maximum of the 173 observations for each variable used in this research.
### Table 1 - Variable Summaries

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Standard Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Federal Funds</strong></td>
<td>2.81</td>
<td>2.51</td>
<td>2.18</td>
<td>0.11</td>
<td>5.33</td>
</tr>
<tr>
<td><strong>Open</strong></td>
<td>1,237.47</td>
<td>1,314.78</td>
<td>237.99</td>
<td>680.76</td>
<td>1,562.25</td>
</tr>
<tr>
<td><strong>High</strong></td>
<td>1,246.76</td>
<td>1,323.78</td>
<td>236.11</td>
<td>695.27</td>
<td>1,564.74</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>1,226.93</td>
<td>1,311.58</td>
<td>241.92</td>
<td>672.88</td>
<td>1,549.00</td>
</tr>
<tr>
<td><strong>Close</strong></td>
<td>1,236.51</td>
<td>1,318.00</td>
<td>239.98</td>
<td>676.53</td>
<td>1,552.58</td>
</tr>
<tr>
<td><strong>Volume</strong></td>
<td>4.03E+09</td>
<td>3.71E+09</td>
<td>1.53E+09</td>
<td>1.27E+09</td>
<td>8.28E+09</td>
</tr>
<tr>
<td><strong>Open1</strong></td>
<td>-0.09</td>
<td>0.05</td>
<td>1.87</td>
<td>-7.95</td>
<td>10.65</td>
</tr>
<tr>
<td><strong>High1</strong></td>
<td>0.00</td>
<td>-0.01</td>
<td>1.25</td>
<td>-4.28</td>
<td>5.23</td>
</tr>
<tr>
<td><strong>Low1</strong></td>
<td>-0.01</td>
<td>0.00</td>
<td>1.23</td>
<td>-4.49</td>
<td>6.50</td>
</tr>
<tr>
<td><strong>Close1</strong></td>
<td>0.19</td>
<td>0.10</td>
<td>1.86</td>
<td>-5.74</td>
<td>10.79</td>
</tr>
<tr>
<td><strong>Prevop1</strong></td>
<td>-0.10</td>
<td>0.00</td>
<td>1.87</td>
<td>-7.95</td>
<td>10.65</td>
</tr>
<tr>
<td><strong>Prevhi1</strong></td>
<td>0.00</td>
<td>-0.01</td>
<td>1.25</td>
<td>-4.28</td>
<td>5.23</td>
</tr>
<tr>
<td><strong>Prevlo1</strong></td>
<td>-0.01</td>
<td>0.00</td>
<td>1.22</td>
<td>-4.49</td>
<td>6.50</td>
</tr>
<tr>
<td><strong>Prevclo1</strong></td>
<td>0.19</td>
<td>0.10</td>
<td>1.86</td>
<td>-5.74</td>
<td>10.79</td>
</tr>
<tr>
<td><strong>Open5</strong></td>
<td>-0.07</td>
<td>0.14</td>
<td>3.34</td>
<td>-16.84</td>
<td>11.47</td>
</tr>
<tr>
<td><strong>High5</strong></td>
<td>-0.05</td>
<td>0.26</td>
<td>2.96</td>
<td>-9.30</td>
<td>11.40</td>
</tr>
<tr>
<td><strong>Low5</strong></td>
<td>0.00</td>
<td>0.14</td>
<td>3.59</td>
<td>-10.26</td>
<td>13.24</td>
</tr>
<tr>
<td><strong>Close5</strong></td>
<td>0.00</td>
<td>0.13</td>
<td>3.52</td>
<td>-13.85</td>
<td>13.83</td>
</tr>
<tr>
<td><strong>Prevop5</strong></td>
<td>-0.08</td>
<td>0.11</td>
<td>3.34</td>
<td>-16.84</td>
<td>11.47</td>
</tr>
<tr>
<td><strong>Prevhi5</strong></td>
<td>-0.06</td>
<td>0.21</td>
<td>2.96</td>
<td>-9.30</td>
<td>11.40</td>
</tr>
<tr>
<td><strong>Prevlo5</strong></td>
<td>-0.02</td>
<td>0.11</td>
<td>3.59</td>
<td>-10.26</td>
<td>13.24</td>
</tr>
<tr>
<td><strong>Prevclo5</strong></td>
<td>-0.01</td>
<td>0.12</td>
<td>3.52</td>
<td>-13.85</td>
<td>13.83</td>
</tr>
<tr>
<td><strong>Open10</strong></td>
<td>-0.25</td>
<td>0.50</td>
<td>4.73</td>
<td>-24.51</td>
<td>13.45</td>
</tr>
<tr>
<td><strong>High10</strong></td>
<td>-0.17</td>
<td>0.52</td>
<td>4.26</td>
<td>-16.72</td>
<td>18.42</td>
</tr>
<tr>
<td><strong>Low10</strong></td>
<td>-0.17</td>
<td>0.43</td>
<td>4.69</td>
<td>-17.50</td>
<td>14.78</td>
</tr>
<tr>
<td><strong>Close10</strong></td>
<td>-0.08</td>
<td>0.42</td>
<td>4.54</td>
<td>-15.39</td>
<td>21.64</td>
</tr>
<tr>
<td><strong>Prevop10</strong></td>
<td>-0.26</td>
<td>0.50</td>
<td>4.73</td>
<td>-24.51</td>
<td>13.45</td>
</tr>
<tr>
<td><strong>Prevhi10</strong></td>
<td>-0.17</td>
<td>0.52</td>
<td>4.26</td>
<td>-16.72</td>
<td>18.42</td>
</tr>
<tr>
<td><strong>Prevlo10</strong></td>
<td>-0.18</td>
<td>0.43</td>
<td>4.69</td>
<td>-17.50</td>
<td>14.78</td>
</tr>
<tr>
<td><strong>Prevclo10</strong></td>
<td>-0.08</td>
<td>0.42</td>
<td>4.54</td>
<td>-15.39</td>
<td>21.64</td>
</tr>
<tr>
<td><strong>Open20</strong></td>
<td>-0.36</td>
<td>1.40</td>
<td>6.58</td>
<td>-27.69</td>
<td>23.35</td>
</tr>
<tr>
<td><strong>High20</strong></td>
<td>-0.19</td>
<td>1.27</td>
<td>5.90</td>
<td>-26.08</td>
<td>20.78</td>
</tr>
<tr>
<td><strong>Low20</strong></td>
<td>-0.24</td>
<td>1.07</td>
<td>6.42</td>
<td>-23.47</td>
<td>22.28</td>
</tr>
<tr>
<td><strong>Close20</strong></td>
<td>-0.02</td>
<td>1.04</td>
<td>6.07</td>
<td>-22.27</td>
<td>23.49</td>
</tr>
<tr>
<td><strong>Prevop20</strong></td>
<td>-0.37</td>
<td>1.34</td>
<td>6.58</td>
<td>-27.69</td>
<td>23.35</td>
</tr>
<tr>
<td><strong>Prevhi20</strong></td>
<td>-0.20</td>
<td>1.27</td>
<td>5.90</td>
<td>-26.08</td>
<td>20.78</td>
</tr>
<tr>
<td><strong>Prevlo20</strong></td>
<td>-0.23</td>
<td>1.15</td>
<td>6.42</td>
<td>-23.47</td>
<td>22.28</td>
</tr>
<tr>
<td><strong>Prevclo20</strong></td>
<td>-0.02</td>
<td>1.06</td>
<td>6.07</td>
<td>-23.27</td>
<td>23.49</td>
</tr>
<tr>
<td><strong>Open60</strong></td>
<td>-0.56</td>
<td>1.79</td>
<td>11.21</td>
<td>-34.54</td>
<td>38.45</td>
</tr>
<tr>
<td><strong>High60</strong></td>
<td>-0.54</td>
<td>1.93</td>
<td>10.82</td>
<td>-33.31</td>
<td>35.56</td>
</tr>
<tr>
<td><strong>Low60</strong></td>
<td>-0.47</td>
<td>1.68</td>
<td>11.35</td>
<td>-34.63</td>
<td>37.30</td>
</tr>
<tr>
<td><strong>Close60</strong></td>
<td>-0.42</td>
<td>2.00</td>
<td>10.97</td>
<td>-33.05</td>
<td>37.73</td>
</tr>
<tr>
<td><strong>Volume1</strong></td>
<td>7.99</td>
<td>8.13</td>
<td>16.03</td>
<td>-73.27</td>
<td>58.63</td>
</tr>
<tr>
<td><strong>Volume5</strong></td>
<td>4.87</td>
<td>3.65</td>
<td>27.59</td>
<td>-64.49</td>
<td>172.42</td>
</tr>
<tr>
<td><strong>Volume10</strong></td>
<td>8.26</td>
<td>2.98</td>
<td>37.04</td>
<td>-56.47</td>
<td>322.20</td>
</tr>
<tr>
<td><strong>Volume20</strong></td>
<td>12.90</td>
<td>7.59</td>
<td>41.31</td>
<td>-65.81</td>
<td>352.52</td>
</tr>
<tr>
<td><strong>Volume60</strong></td>
<td>18.86</td>
<td>12.71</td>
<td>37.89</td>
<td>-47.55</td>
<td>254.97</td>
</tr>
<tr>
<td><strong>Bearish</strong></td>
<td>35.09</td>
<td>35.00</td>
<td>10.20</td>
<td>0.00</td>
<td>65.22</td>
</tr>
<tr>
<td><strong>Neutral</strong></td>
<td>27.31</td>
<td>26.67</td>
<td>9.34</td>
<td>5.00</td>
<td>55.56</td>
</tr>
<tr>
<td><strong>Bullish</strong></td>
<td>37.60</td>
<td>35.30</td>
<td>11.74</td>
<td>13.15</td>
<td>75.00</td>
</tr>
<tr>
<td><strong>Diff</strong></td>
<td>45.39</td>
<td>46.66</td>
<td>18.69</td>
<td>-11.12</td>
<td>90.00</td>
</tr>
</tbody>
</table>
Figure 2 shows closing prices for the S&P 500 for each poll date during our time period as well as the bullish poll percentages less the bearish poll percentages (to act as a measure of blogger bullishness). As you can see, there seems to be little correlation between the changes in blogger sentiment and market prices over time. Sentiment swings quite wildly compared to the market prices, which may take away from any predictive power it might have. An initial look at this data gives the idea that no correlation exists between blogger sentiment and market prices.
5 Results and Analysis

Market Price Regressions

Linear Model

The first set of regressions are simple linear regressions aimed at exploring whether increases in bullish, bearish or neutral percentages in the Blogger Sentiment Poll are correlated with percentage changes in market price growth rates. Therefore, the dependent variables (one for each regression) are the 1-day, 5-day, 10-day and 20-day growth rates for the opening, high, low, and closing market prices. The independent variables are first one out of the three bullish, bearish or neutral percentages and the federal funds rate, and then two of the three poll percentages together (all three together always add up to 100, so a maximum of two must be used), and the federal funds rate. All regressions were tested and corrected for any heteroskedasticity or autocorrelation using the “robust” command in Stata as well as the Cochrane-Orcutt algorithm.

Results for the 1-day and 5-day growth rates were generally not significant (all regressions are tested at the 10% significance level), and when they were, did not follow any sort of pattern. This makes sense, since short-term (1-day, 1-week) market movements are usually more volatile and less cohesive than longer-term movements.

Results for the 10-day, 20-day and 60-day price growth rates are very similar in their significance and coefficients: when the bullish variable is significant its coefficient on growth rates is negative; when the bearish variable is significant its coefficient is positive; and when the neutral variable is significant its coefficient is negative. The bullish and bearish variables were significant in almost all of these regressions, while the neutral variable was generally only significant in regressions run with it as the only poll variable or alongside the bearish variable.
Linear Model Analysis

Clearly, this is an unexpected find – if the bloggers were generally correct, we would see bullish variables with a positive coefficient on growth rates, bearish variables with a negative coefficient, and neutral variables with insignificant coefficients. This format would mean that increases in the percent of bloggers who predict the market to be bullish would be correlated with an increase in growth rates, increases in the percent of bloggers who predict the market to be bearish would correlated with a decrease in growth rates, and an increase in the percent of bloggers who predict the market to be neutral would have no measurable effect on the growth rates, which is what you would expect if the bloggers were correct.

However, the averages of the significant results say that a 1% increase in the number of bullish bloggers is correlated with a 0.0829646 decrease in the 2-week growth rate, a .090002 decrease in the 1-month growth rate, and a 0.09137 decrease in the 3-month growth rate; a 1% increase in the number of bearish bloggers is correlated with a 0.103956 increase in the 2-week growth rate, a 0.104676 increase in the 1-month growth rate, and a 0.101631 increase in the 3-month growth rate; a 1% increase in the number of neutral bloggers is correlated with a 0.110499 decrease in the 2-week growth rate, a 0.102473 decrease in the 1-month growth rate, and a 0.095748 decrease in the 3-month growth rate.

Probit/Logit Models

That these significant coefficients were the opposite of what one might expect – basically pointing out that these well-respected financial bloggers are mostly wrong in their predictions – warranted further inspection. The next series of regressions used the dummy variables created to signify when a particular opening, high, low or closing growth rate fell into a bearish, bullish or neutral range for that particular growth rate series (as explained in the Data section). Probit and
logit models allow these dummy variables (such as “high5bu,” the dummy variable equaling 1 when particular values of the high 5-day market growth rate fall into the bullish range of that data and 0 when they do not) to be dependent variables, so the same sequence of independent variables as in the previous linear regressions can go on the right hand side of the regression equations. This gives greater insight as to what the bullish, bearish and neutral blogger predictions are really telling us about market price movements.

Once again, the 1-day and 5-day growth rates were generally not significant, and when they were, did not offer a very cohesive pattern. For the 10-day growth rates, the only significant pattern that shows up is for our control variable, the federal funds rate. These regressions show that increases in the federal funds rate increase the probability of the growth rate over the next ten days being neutral.

20-day and 60-day results were similar in the signs of the coefficients, with the 60-day results having even lower p-values, stronger coefficients and better patterns than the 20-day results. The general patterns in these bullish, bearish and neutral growth rate regressions are as follows: for bullish growth rates, the bullish poll percentage and federal funds variables have negative coefficients, while the bearish poll percentage variable has a positive coefficient. For bearish growth rates, the bullish poll percentage has a positive coefficient. Finally, for the neutral growth rates, both the neutral poll percentage and federal funds variables have positive coefficients.
Tables 2 and 3 show these significant relationships found between the independent variables (the bullish, bearish and neutral percentages of the blogger sentiment poll along with the federal funds rate) and the dependent variables (dummy variables signifying bullish, bearish and neutral one-month and three-month growth rates for opening, high, low and closing market
prices). Six sets of regressions were run for each of these dependent dummy variables with different combinations of bullish, bearish and neutral poll percentages as the independent variables. Positive signs indicate that an independent variable increases the probability of that particular opening, high, low or closing market growth rate being bullish, bearish or neutral, while negative signs indicate decreases in that probability. Blanks cells denote a lack of significance at the 10% level.

**Probit/Logit Model Analysis**

Since the results of the probit and logit models were similar, for the sake of interpretation the probit models (using the “dprobit” command in Stata) will be used in further analysis. Again, we are seeing more of the backwards results that we found in the linear regressions, but the results from these new regressions give us real advice as to what changes in the blogger market outlook percentages can say about future market growth rates.

For our 10-day growth rates, the probit models tells us that, on average, a 1% increase in the federal funds rate increases the probability of the 2-week growth rate being neutral by 5.5172%.

For the 20-day growth rates, on average a 1% increase in the number of bullish bloggers decreases the probability of the growth rate being bullish by 0.8157%, while a 1% increase in the federal funds rate decreases that same probability by 3.3838%. A 1% increase in the number of bearish bloggers increases this probability by 0.8575%. A 1% increase in the number of bullish bloggers heightens the probability of the growth rate being bearish by 0.8488%. A 1% increase in the number of neutral bloggers raises the probability of the growth rate being neutral by 1.0382%, while a 1% increase in the federal funds rate raises that same probability by 4.1801%.
For the 60-day growth rates, on average a 1% increase in the number of bullish bloggers decreases the probability of the growth rate being bullish by 1.325%, while a 1% increase in the federal funds rate lowers the same probability by 4.3282%. A 1% increase in the number of bearish bloggers raises this probability by 1.1163%. A 1% boost in the number of neutral bloggers decreases the probability of the growth rate being bearish by 1.7061%, while a 1% increase in the number of bullish bloggers heightens this probability by 1.363%. A 1% increase in the number of neutral bloggers improves the probability of the growth rate being neutral by 0.9249%, while a 1% increase in the federal funds rate strengthens the same probability by 6.345%.

Again, our results from the linear regressions are confirmed and refined. The more bullish the bloggers are the less likely the growth rates are to be bullish and the more likely they are to be bearish. In other words, when these bloggers increase in bullishness, be wary. Increases in bearishness, however, are usually a good sign for the market – higher blogger bearishness increases the likelihood of these growth rates being bullish. The only category the bloggers seem to get right is neutral growth – increases in neutral bloggers generally do point to a more neutral market (and a less bearish one) over the next one-month and three-month periods.

One interesting thing to note in these regressions is the federal funds rate – when it is significant, it has a much greater impact on the bullishness, bearishness or neutrality of the growth rates than do the poll percentages. Having an increase in the federal funds rate decreasing the probability of the market being bullish makes sense – increases in the federal funds rate raise the level at which investors are discounting future cash flows, making future dividends and stock growths less valuable at the present time, which would then be reflected in the markets through lowered current prices. However, you would then expect to have an increase in the federal funds
rate heighten the likelihood of a bearish market, which we do not see with any significance. Also, it is peculiar that increases in the funds rate would make a neutral market more likely (having an even greater effect than on the likelihood of a bullish market). This could be the case due to the interesting time period over which this data is observed – a period featuring great volatility with historically low interest rates. However, the bottom line is that, when trying to predict future market price growth rates, this data says that you would be much better off using the federal funds rate as your guide than the word of these financial bloggers.

**Time-lag Regressions**

Because of the seemingly glaring inaccuracies of these well-respected financial bloggers, I ran a series of linear regressions to explore what these bloggers were basing their predictions on – most importantly, whether they were basing their future predictions on past market performance, a common investing error. I created time-lag data, variables that gave the previous day, week, two-week, and month-long growth rates of the opening, high, low and closing market values for each date the poll was released. Because the bloggers submit their predictions over the weekend, even the previous day’s growth rate (each Friday) could have an effect. So, the dependent variables in these regressions were one of the bullish, bearish and neutral percentages of the poll, and the independent variables were one of these previous growth rates along with the federal funds rate.

The previous day values were generally significant for both bullish and bearish poll percentages – 1% increases in the previous day’s growth rate are correlated with an average increase in the percentage of bullish bloggers of 1.34574 and an average decrease in the percentage of bearish bloggers of 1.38065. Only the number of bearish bloggers was affected by the growth rate of the previous week – again with a negative coefficient (this time an average of -
Previous two-week and month-long growth rates had no significant effect on the percentage of bullish or bearish bloggers. Interestingly enough, the previous week, two-week, and month-long growth rates all had a positive correlation with the percentage of neutral bloggers – a 1% increase in these previous growth rates correlated with average increases in the percentage of neutral bloggers of 0.332747. Finally, the federal funds rate was significantly correlated with both bearish and bullish poll percentages in each regression in just the way you would expect – negatively for bullish, and positively for bearish.

**Analysis**

This information points to the possibility that the market outlook these bloggers give in the poll is affected by the previous day’s performance, but is only affected by the previous week’s performance if that performance was negative. The positive correlation that the previous week, two-week and month-long growth rates had with neutral predictions might be explained again by the great volatility and uncertainty during much of the time period this data is from, and the fact that a majority of the bloggers were bearish in their predictions for much of this time. Recent positive performance may have caused bloggers to change their outlook to neutral, but wasn’t enough to increase their month-long outlook to bullish.

The most important part of these regressions is, again, the significance of the coefficients on the federal funds rate, acting as our bellwether for economic health and market movements. The coefficients on the rate were similar for all bullish regressions, all bearish regressions, and all neutral regressions over the different growth rates. On average, a 1% increase in the federal funds rate was correlated with a 1.44189 decrease in the percentage of bullish bloggers and a 1.06652 increase in the percentage of bearish bloggers.
The fact that the rate acted in just the way one would expect, and that its significant coefficients were much greater in magnitude than those of the lag growth variables, gives power to the idea that the bloggers were basing their predictions off of general market information that they had at the time much more so than past market performance.

**Volume Regressions**

This final section of regressions explores whether changes in the blogger sentiment are correlated with movements in the market trading volume. Volume is the number of trades that are processed in the stock market each day. Theoretically, one would expect that investor outlook on the market would have a lot to do with the level of trading – more specifically, I wanted to test the theory that heightened disagreement in market outlook between investors would cause increases in trading volume. Intuitively this makes sense – investors bullish on the market would look to buy stocks, and investors bearish on the market would look to sell them. If the levels of both bearish and bullish investors rose at the same time (obviously with decreases in the number of neutral investors), there would be more sellers for every buyer and vice versa, hence more trades would happen and trade volume would rise. Using the results of the blogger sentiment poll as a measure of broader investment sentiment makes this analyzation broader in scope. For the first set of regressions the dependent variables were the 1-day, 5-day, 10-day, 20-day and 60-day volume growth rates, with the independent variables being either the “Diff” variable (a measure of difference in blogger sentiment, see the Data section) or two out of the three bullish, bearish or neutral blog percentage values, both along with the federal funds rate as the market control variable.
Analysis

The coefficient on the difference variable was only significant for the 60-day volume growth rate, showing little relationship between difference in market forecasting opinion (or investor sentiment) and the trade volume, similar to what was found in Antweiler and Frank [2001]. Bullish, bearish and neutral variables also had few significant coefficients (almost solely on the 60-day volume growth rates) and thus show a similar lack of correlation between whether market participants are bullish, bearish or neutral and the magnitude of the trade volume. The only repeated significant coefficients were on the federal funds rate variable: a 1% increase in the federal funds rate on average correlated to a 1.19138% rise in the 1-day volume growth rate, a 2.77266% rise in the month-long volume growth rate, and a 3.80159% rise in the 3-month volume growth rate. The significance and pattern of these coefficients suggest that changes in trade volume depends much more so on general market conditions, which in our case are symbolized by the federal funds rate, than on investor sentiment.
6 Conclusion

In this paper I analyzed the accuracy of market predictions made by financial bloggers, focusing on what their forecasts and sentiment could tell us about future market price and volume movements, and looked at what variables affected their sentiment in the first place. I found that these well-respected bloggers were mostly wrong in their predictions – most notably that greater bullish sentiment within the bloggers generally led to lower future growth rates, and increases in bearish sentiment generally led to higher future growth rates. In other words, investor sentiment is a weak indicator for future performance. The only predictions that had any level of accuracy were the neutral market predictions.

In terms of changes in volume growth rates, I found that increases in sentiment difference between the bloggers did not lead to increases in volume as was initially hypothesized. Increases in bullishness and bearishness of the bloggers had little effect on volume changes as well.

Looking at whether or not the bloggers were fooled into inaccurate predictions by basing their guesses on recent market performance, results showed that bloggers were affected by the previous day’s performance, as well as the previous week’s performance if that performance was bearish.

In all of my regressions I found that the federal funds rate, in this case used to control for overall macroeconomic conditions, had correlation with the dependent variables much greater in magnitude than did any other independent variables. Increases in the federal funds rate indicated decreases in the probability of market growth rates being bullish (as one would expect), as well as increases in the probability of market growth rates being neutral (unexpected). This latter find may be due to the long period of historically low interest rates during the great recession of 2008.
and 2009 – rises in interest rates coming out of this recession only occurred when the market was deemed stable enough to handle less liquidity.

The federal funds rate also had greater effects on changes in volume than did blogger sentiment. Increases in the federal funds rate had significant correlation with rises in volume growth rates over one day, one week, two week, one month and three month periods. This suggests that the level of trading that occurs is affected more so by overall market conditions and macroeconomic policy than investor sentiment.

Finally, the federal funds rate had a significant effect on the predictions made by the financial bloggers. Increases in the federal funds rate led to lower percentages of bullish bloggers and higher percentages of bearish bloggers, reactions that make theoretical sense due to the rate’s affect on discounting future cash flows. The significance and magnitude of these coefficients indicate that the bloggers are generally basing their predictions off of present economic and financial indicators, as would be expected.

Further exploration of this material could include adding more economic indicator variables to get a better idea of what is affecting the sentiment of these bloggers. Extreme points of blogger bullishness or bearishness could be analyzed. Finally, sets of blog posts could be rated on a bullish/bearish scale using linguistics programs to see if the orientation of the posts has a day-to-day effect on market movements, effectively investigating the power that these bloggers have over the market.
7 Appendix

Example 1 – Significant output from a linear regression on market price

```
.prais close20 bullish neutral federal_funds, corc
Iteration 0:  rho = 0.0000
Iteration 1:  rho = 0.7292
Iteration 2:  rho = 0.7530
Iteration 3:  rho = 0.7536
Iteration 4:  rho = 0.7536
Iteration 5:  rho = 0.7536

Cochrane-Orcutt AR(1) regression -- iterated estimates
```

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 172</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>269.28079</td>
<td>3</td>
<td>89.7602635</td>
<td>F(3, 168) = 5.58</td>
</tr>
<tr>
<td>Residual</td>
<td>2702.25593</td>
<td>168</td>
<td>16.0848567</td>
<td>Prob &gt; F = 0.0011</td>
</tr>
<tr>
<td>Total</td>
<td>2971.53672</td>
<td>171</td>
<td>17.3774077</td>
<td>R-squared = 0.0906</td>
</tr>
</tbody>
</table>

|             | Coef.       | Std. Err. | t    | P>|t|   | [95% Conf. Interval] |
|-------------|-------------|-----------|------|-------|---------------------|
| close20     |             |           |      |       |                     |
| bullish     | -.1222873   | .037412   | -3.27| 0.001 | -.1961455 -.048429  |
| neutral     | -.1422707   | .038842   | -3.66| 0.000 | -.218952 -.0655895  |
| federal_funds | .0082646   | .5444663  | 0.02 | 0.813 | -1.06613 1.083142   |
| _cons       | 8.410421    | 2.97936   | 2.82 | 0.005 | 2.528612 14.29223   |
| rho         | .7536396    |           |      |       |                     |

Durbin-Watson statistic (original) 0.541473
Durbin-Watson statistic (transformed) 2.025108

Example 2 – Significant output from a probit regression on market price

```
dprobit close60bu bullish neutral federal_funds
Iteration 0:  log likelihood = -113.44912
Iteration 1:  log likelihood = -106.04513
Iteration 2:  log likelihood = -105.98482
Iteration 3:  log likelihood = -105.98481

Probit regression, reporting marginal effects  Number of obs = 173
LR chi2(3) = 14.93
Prob > chi2 = 0.0019
Pseudo R2 = 0.0658
```

| close6~u  | dF/dx     | Std. Err. | z     | P>|z| | x-bar [    95% C.I.   ] |
|-----------|-----------|-----------|-------|------|------------------------|
| bullish   | -.013021  | .0041941  | -3.08 | 0.002| 37.5994 -.021241 -.004801|
| neutral   | -.0011205 | .0047482  | -0.24 | 0.813| 27.3072 -.010427 .008186|
| federal~s | -.0403868 | .0177029  | -2.28 | 0.022| 2.81393 -0.75084 -0.00569|
| obs. P    | .3641618  |           |       |      |                        |
| pred. P   | .3536361  | (at x-bar)|      |      |                        |

z and P>|z| correspond to the test of the underlying coefficient being 0
Example 3 – Significant output from a linear regression on volume

```
. regress vol60 diff federal_funds

Source |       SS       df       MS              Number of obs =     173
-------------+------------------------------           F(  2,   170) =    4.74
Model |  13050.4795     2  6525.23975           Prob > F      =  0.0099
  Residual |  233893.965   170  1375.84685           R-squared     =  0.0528
  Total |  246944.444   172  1435.72351           Adj R-squared =  0.0417

------------------------------------------------------------------------------
vol60 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
diff |   .3081438   .1519921     2.03   0.044     .0081088    .6081788
federal_funds |    3.24388   1.299991     2.50   0.014     .6776747    5.810085
_cons |  -4.255973    8.57721    -0.50   0.620    -21.18753    12.67558
------------------------------------------------------------------------------
```

Example 4 – Significant output from a linear regression on blogger sentiment

```
. regress bullish prevclo1 federal_funds

Source |       SS       df       MS              Number of obs =     173
-------------+------------------------------           F(  2,   170) =   10.07
Model |  2511.78214     2  1255.89107           Prob > F      =  0.0001
  Residual |   21192.274   170  124.660436           R-squared     =  0.1060
  Total |  23704.0562   172   137.81428           Adj R-squared =  0.0954

------------------------------------------------------------------------------
bullish |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
prevclo1 |   1.137343   .4591427     2.48   0.014     .2309875    2.043698
federal_funds |  -1.356605   .3918153    -3.46   0.001    -2.130055   -.5831551
_cons |   41.20577   1.401131    29.41   0.000     38.43991    43.97162
------------------------------------------------------------------------------
```
8 References


